

Discussion Papers

Deutsches Institut für Wirtschaftsforschung

2012

Disentangling Demand and Supply Shocks in the Crude Oil Market

How to Check Sign Restrictions in Structural VARs

Helmut Lutkepohl and Aleksei Netšunajev

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2012

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

February 10, 2012

Disentangling Demand and Supply Shocks in the Crude Oil Market: How to Check Sign Restrictions in Structural VARs¹

Helmut Lütkepohl

Department of Economics, Freie Universität Berlin and DIW Berlin
Mohrenstr. 58, D-10117 Berlin, Germany
email: hluetkepohl@diw.de

and

Aleksei Netsunajev

Department of Economics, European University Institute
Via della Piazzola 43, I-50133 Firenze, Italy
email: aleksei.netsunajev@eui.eu

Abstract. Given the growing dissatisfaction with exclusion and long-run restrictions in structural vector autoregressive analysis, sign restrictions are becoming increasingly popular. So far there are no techniques for validating the shocks identified via such restrictions. Although in an ideal setting the sign restrictions specify shocks of interest, sign restrictions may be invalidated by measurement errors, data adjustments or omitted variables. We model changes in the volatility of the shocks via a Markov switching (MS) mechanism and use this device to give the data a chance to object to sign restrictions. The approach is illustrated by considering a small model for the market of crude oil.

Key Words: Markov switching model, vector autoregression, heteroskedasticity, crude oil market

JEL classification: C32, Q43

¹We thank Lutz Kilian, Helmut Herwartz and participants of the *EC²* Conference in Florence in December 2011 as well as seminars at the University of Venice, the University of Zurich and the Deutsche Bundesbank for comments on an earlier version of this paper. The research for this paper was carried out in part while the first author was a Bundesbank Professor.

1 Introduction

Identifying the shocks of interest is a major problem in structural vector autoregressive (SVAR) analysis. Often just-identifying restrictions are imposed, for example on the instantaneous effects of the shocks as in Sims (1980) and Amisano and Giannini (1997) or on the long-run effects as in Blanchard and Quah (1989) and King, Plosser, Stock and Watson (1991) (see Lütkepohl (2005) for a textbook exposition). Typically the restrictions imposed thereby are just-identifying and, hence, cannot be checked against the data by statistical tests. There is some dissatisfaction with these types of equality restrictions because there is often no agreement on them. Reasons may be that different underlying economic models imply alternative sets of restrictions or that they may not imply sufficiently precise restrictions to uniquely identify all shocks of interest.

Therefore identification by inequality or sign restrictions for the effects of shocks on certain variables is proposed by a number of authors (see, e.g., Faust (1998), Canova and De Nicoló (2002), Uhlig (2005)). In the latter article shocks are identified by specifying the sign of their effect on certain variables on impact and possibly in some of the following periods. All shocks that are in line with the sign restrictions are considered admissible. Since it appears to be easier to agree on such softer restrictions, this approach is seen as an attractive alternative to identification via more classical exclusion restrictions and consequently it has been applied by a number of researchers (e.g., Mountford and Uhlig (2009), Lippi and Nobili (2011), Peersman and Straub (2009), Canova and Pappa (2007)).

In some literature, dynamic stochastic general equilibrium (DSGE) models are used to check the signs of responses to shocks of interest and the corresponding sign restrictions are used to characterize the shocks and are then imposed on a SVAR model for the identification of the shocks (see Canova (2002), Dedola and Neri (2007), Pappa (2009)). DSGE models are highly stylized models of an economy and as such not necessarily a good description of the data. Therefore it makes sense to assume that they cannot give a precise picture of the reactions to shocks in actual economic systems but may be able to suggest at least a direction of the responses of the variables to shocks. Using these directions or signs of the responses in a VAR model which is specified so as to fit the data well, is hoped to provide a better understanding of the effects of shocks in practice.

Unfortunately, using sign restrictions for identifying structural shocks has some drawbacks as well. First of all, they typically do not identify the shocks uniquely but allow for a range of admissible shocks that are all in line with the sign restrictions. Therefore the range of possible responses may also be

large and, hence, sign restrictions may deliver only a rather diffuse picture of the reactions of variables to a shock of interest. Second, identification via sign restrictions does not provide a possibility to check the validity of the restrictions because only those shocks are considered that satisfy the prespecified sign restrictions. Thus, any analysis based upon them is conditional on the sign restrictions being correct. In other words, the data have no possibility to speak up against the restrictions. This criticism has been raised against conventional just-identifying restrictions as well, of course, but it is no less valid in the context of sign restrictions. One may argue that the problem is less severe in the context of sign restrictions because they are much weaker restrictions than equality restrictions. Even if the sign restrictions are derived from generally accepted economic models, there is a potential gap between the empirical and the theoretical models that may invalidate the sign restrictions in the empirical model. Reasons may be that the variables used in the empirical model do not correspond exactly to those considered in the theoretical model, for instance, due to measurement errors, trend and/or seasonal adjustment or using data with an observation frequency that does not correspond to that of the theoretical model. In addition, the economic model describes the relations within a set of variables and may explicitly abstract from other effects that may, however, be important in practice. For example, the theoretical model may be one for a closed economy whereas foreign effects may not be negligible in the actual system and the available data. In other words, there may be an omitted variables problem that could affect the empirical model.

In this study a proposal is made as to how to let the data speak about the validity of identifying restrictions in general and sign restrictions in particular. The idea is to use changes in volatility to support the identification of shocks. In SVAR analysis identification via heteroskedasticity is proposed and used by Rigobon (2003), Rigobon and Sack (2003), Lanne and Lütkepohl (2008) among others. These authors essentially assume that there are exogenously generated changes in the volatility of the shocks and partition the sample period accordingly. Then they base the identification of the shocks on the assumption that the effects of shocks are the same regardless of the volatility regime in which they occur. In other words, they assume that the impulse responses are invariant throughout the sample period whereas the volatility of the shocks may change. Thereby additional, statistical identifying restrictions become available, which can be used to check restrictions that are just-identifying in the conventional approach. We will use the approach proposed by Lanne, Lütkepohl and Maciejowska (2010) and model the changes in volatility endogenously by means of a Markov regime switching (MS) mechanism. A related proposal, based on different models for condi-

tional heteroskedasticity, is used by Normandin and Phaneuf (2004), Bouakez and Normandin (2010) and others for identifying shocks.

In the present study it will be shown how changes in volatility can provide additional identifying information that can be used for checking the validity of sign restrictions. Of course, this approach requires that there are changes in volatility during the sample period and cannot be used if volatility is time-invariant. If volatility varies sufficiently this device can generate a full set of unique shocks that are in line with the data. These shocks may not have an economic interpretation. If none satisfy the sign restrictions it means that there is no shock that is acceptable to the data and satisfies the sign restrictions at the same time. For example, if there are sign restrictions characterizing a shock as a technology shock and none of the shocks acceptable to the data satisfies those restrictions, one may conclude that a technology shock cannot be isolated in the system under consideration. Using the shocks identified by changes in volatility may result in unique shocks and not in a range of feasible shocks and, hence, it may also provide more precise impulse responses. In any case, the additional identification information obtained from heteroskedasticity or conditional heteroskedasticity can be helpful for checking the validity of the sign restrictions in the empirical model. Consequently, the two drawbacks of sign restrictions mentioned earlier, that is, lack of precise impulse responses and the inability to reject the restrictions, can be addressed if there are changes in volatility.

One may argue that the typical approach for sign restricted SVARs uses Bayesian methods for estimation that are based on independently, identically distributed (i.i.d.) residuals. Hence, changes in volatility are excluded and our model framework does not nest the typical model assumed under sign restrictions. Although such arguments are valid, it may be worth noting that sign restrictions can be used in conjunction with classical estimation methods (see Moon and Schorfheide (2009) and Moon, Schorfheide, Granziera and Lee (2009) for a discussion of the frequentist approach to inference in such models). In that framework they can be justified under assumptions more general than i.i.d. for the errors. Moreover, even though Bayesian methods require stringent assumptions regarding the error distribution of the VAR model, in applied work they will typically be approximately satisfied at best. Our approach also provides a statistical check of some of these assumptions.

In this study we use a classical approach for inference. In other words, we use likelihood based methods to avoid any distortions by imposing priors on the parameters although Bayesian methods are more common in conjunction with sign restrictions. However, our main arguments do not depend on the method for estimation. They could be equally well placed in a Bayesian framework.

To illustrate the approach, a study by Kilian and Murphy (2011) will be reconsidered. These authors analyze a system with the following three variables:

- $\Delta prod_t$ - percent change in global crude oil production,
- q_t - log detrended index of real economic activity,
- p_t - log of real price of oil.

They specify an oil supply shock, an aggregate demand shock and an oil-market specific demand shock purely by sign restrictions. They point out that sign restrictions are not enough to get a precise picture of the effects of such shocks and they propose to add further information to pin down the effects more precisely. The example is particularly suitable for our purposes because changes in volatility of oil production and the price of oil are diagnosed in the related literature (see Baumeister and Peersman (2010)). In our framework these volatility changes can be used to address the question whether the system is suitable for isolating the three shocks of interest and also to obtain more information which can help to narrow down the effects of the shocks.

The study is organized as follows. The model setup and some technical details of our inference procedures are given in Section 2. The empirical study illustrating the method for checking the sign restrictions is presented in Section 3 and conclusions are provided in Section 4.

2 The Model Setup and Inference

The reduced form of our model is a K -dimensional VAR(p),

$$Y_t = \nu + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + U_t, \quad (1)$$

where ν is a constant term, the A_j ($j = 1, \dots, p$) are $(K \times K)$ coefficient matrices and U_t is a zero-mean error term.

In a conventional SVAR model the structural shocks are obtained from the reduced form residuals by a linear transformation, $\varepsilon_t = B^{-1}U_t$ or $B\varepsilon_t = U_t$, where B is such that ε_t has identity covariance matrix, that is, $\varepsilon_t \sim (0, I_K)$, and the reduced form residual covariance matrix is decomposed as $E(U_t U_t') = \Sigma_U = BB'$.

Following Lanne et al. (2010), in our setup the distribution of the reduced form error term U_t is assumed to depend on a discrete Markov process s_t ($t = 0, \pm 1, \pm 2, \dots$) with states $1, \dots, M$ and transition probabilities

$$p_{ij} = \Pr(s_t = j | s_{t-1} = i), \quad i, j = 1, \dots, M.$$

The conditional distribution of U_t given s_t is assumed to be normal,

$$U_t|s_t \sim \mathcal{N}(0, \Sigma_{s_t}). \quad (2)$$

Our model allows for Markov switching in the residual covariances only and not in other parameters of the model. Although there are finitely many states only and in practice the number of different states will be small, the model can capture smooth transitions from one state to another because a particular state does not necessarily come up with probability one but the system may be in between states, that is, in a given period the actual volatility may be described by a mixture of different states, each state being weighted by a certain probability. Thus, in this sense our model is similar to a multivariate GARCH or stochastic volatility model and can capture similar changes in volatility. We do not allow other parameters than the residual covariance to be state-dependent. Thereby we impose more regularity on our models than in the MS-SVAR models considered by Rubio-Ramirez, Waggoner and Zha (2005), Sims and Zha (2006) and Sims, Waggoner and Zha (2008) or in SVAR models with time-varying coefficients (see, e.g., Cogley and Sargent (2005), Primiceri (2005), Baumeister and Peersman (2010)). In the latter paper, a system similar to our example system analyzed in Section 3 is considered in the framework of a time-varying coefficient SVAR model. We will return to their results in the empirical section.

In the following we will use the abbreviation MS(M)-VAR(p) model for a VAR model with p lags and M Markov states. We emphasize that we use this notation for simplicity although all VAR coefficients apart from the residual covariances are state-invariant. This fact is not reflected in our notation. For example, Krolzig (1997) uses the notation MSH-VAR for our model type.

The fact that the covariances Σ_{s_t} can vary across states is used in our framework for identifying structural shocks that are consistent with the statistical properties of the data. For example, if there are just two states ($M = 2$), then there exists a decomposition $\Sigma_1 = BB'$ and $\Sigma_2 = B\Lambda_2B'$, where $\Lambda_2 = \text{diag}(\lambda_{21}, \dots, \lambda_{2K})$ is a diagonal matrix with positive diagonal elements. If the λ_{2i} s are all distinct, this decomposition is unique apart from changes in the sign and permutations of the columns of B and corresponding changes in the ordering of the λ_{2i} s (Lanne et al. (2010)). Thus, for a given ordering of the λ_{2k} , if the structural shocks are orthogonal, have the same instantaneous effects across states and are normalized such that they have unit conditional variance in the first state, then they are uniquely determined by the transformation $\varepsilon_t = B^{-1}U_t$. Thus, any additional restrictions from economics become over-identifying and can be checked against the data. For instance, exclusion restrictions can be checked with formal statistical tests.

Moreover, if a set of sign restrictions is postulated for the system to be suitable for the desired analysis, the impulse responses corresponding to the shocks identified via the MS structure must satisfy the prespecified sign restrictions that characterize the economic shocks of interest. Alternatively, if the shocks do not satisfy the sign restrictions, these restrictions are not compatible with the statistical properties of the data. The reasons could be those mentioned in the introduction, that is, omitted variables, measurement errors, aggregation problems or distortions due to data transformations. If, however, the impulse responses satisfy the sign restrictions, labels may be attached to the shocks accordingly.

It may be worth emphasizing that the requirement of having distinct λ_{2i} s is crucial for exact identification of all shocks. The λ_{2i} s can be interpreted as variances in State 2 relative to those in State 1. Thus, distinct λ_{2i} s imply that the volatility changes are not homogenous across all variables. Moreover, one of the λ_{2i} s may, of course, be 1, that is, one of the shocks may have the same variance in both states, as long as there is enough heterogeneity in the volatility of the other shocks. Moreover, even if some of the λ_{2i} s are identical, it may still be possible to identify those shocks with distinct relative variances in State 2. Of course, in that situation it may not be possible to arrive at clear conclusions regarding the validity of the sign restrictions. An important advantage of our approach is that the crucial identification restrictions (distinct λ_{2i} s) can be checked with statistical tests rather than having to assume it.

If there are more than two volatility states, the corresponding covariance matrix decomposition

$$\Sigma_1 = BB', \quad \Sigma_i = B\Lambda_i B', \quad i = 2, \dots, M, \quad (3)$$

with diagonal Λ_i matrices is restrictive. Lanne et al. (2010) discuss a likelihood ratio (LR) test for these restrictions. Denoting the diagonal elements of Λ_j by $\lambda_{j1}, \dots, \lambda_{jK}$, uniqueness of B up to sign and permutation of the shocks is ensured when there are more than two states if for any subscripts $k, l \in \{1, \dots, K\}$, $k \neq l$, there is a $j \in \{2, \dots, M\}$ such that $\lambda_{jk} \neq \lambda_{jl}$ (Lanne et al. (2010, Proposition 1)). Although this condition for exact (local) identification is apparently more complicated than in the 2-state case, it is potentially also more likely to be satisfied if there are more than two states and, hence, more heterogeneity in the volatility. Also in this case all shocks may be identified even if one of them has homogenous volatility across all states. Again, the identification condition can be checked by statistical tests. If it is satisfied, the resulting shocks are unique and need to satisfy any valid sign restrictions. In turn, any restrictions that the shocks do not satisfy, are not compatible with the data.

Our identification techniques are the same as those used in the literature on identification via heteroskedasticity (see, in particular, Lanne and Lütkepohl (2008)). While in that literature heteroskedasticity is used, we consider conditional heteroskedasticity. In our approach changes in volatility are determined endogenously from the data. In principle, the periods associated with a particular volatility state can be spread irregularly throughout the sample. Thus, the volatility states are more flexible than in the standard approach based on heteroskedasticity.

Since we assume normality of the residuals conditional on the states, the likelihood function can be set up and the model can be estimated by maximum likelihood (ML). The likelihood function is given in Lanne et al. (2010) and a detailed discussion of the related estimation problems can be found in Herwartz and Lütkepohl (2011). They also present the details of an EM algorithm for optimizing the likelihood and point out that there are many local maxima. In fact, the numerical problems are challenging, in particular when a large number of different Markov states are allowed for. The procedure is a quasi ML procedure if the normality assumption in (2) does not hold. The normality assumption is not essential for the asymptotic properties of the estimates but is just used for setting up the likelihood function.

Herwartz and Lütkepohl (2011) discuss a fixed design wild bootstrap procedure for constructing confidence intervals for impulse responses in the presently considered model class. They propose to construct bootstrap samples conditional on the ML estimates so that for our model setup they are constructed as

$$Y_t^* = \hat{\nu} + \hat{A}_1 Y_{t-1} + \dots + \hat{A}_p Y_{t-p} + U_t^*, \quad (4)$$

where $U_t^* = \eta_t \hat{U}_t$ and η_t is a random variable with values 1 and -1 , each with probability 0.5. Thereby potential heteroskedasticity and the pattern of contemporaneous dependence of the data is preserved. We bootstrap parameter estimates θ^* of $\theta = \text{vec}[\nu, A_1, \dots, A_p]$ and B^* of B , conditionally on the initially estimated transition probabilities and Λ_m , $m = 2, \dots, M$, to alleviate the computational burden as in Herwartz and Lütkepohl (2011). Notice that computing the bootstrap impulse responses still requires nonlinear optimization of the log-likelihood and, hence, is computationally demanding. We use the ML estimates as starting values in the bootstrap replications. In our empirical analysis we consider 68% standard percentile confidence intervals based on 1000 replications. The confidence level is in line with much of the related literature.

3 Empirical Analysis

3.1 Previous identification restrictions

Kilian and Murphy (2011) are primarily interested in the effects of demand and supply shocks in the crude oil market on the real price of oil. They argue that the exclusion restrictions used by Kilian (2009) for identifying the shocks could be questioned. In the latter article Kilian identifies the shocks by assuming that oil-market specific demand shocks ($\varepsilon_t^{\text{oil-d}}$) do not have an instantaneous effect on oil production and real activity and aggregate demand shocks ($\varepsilon_t^{\text{aggr-d}}$) have no immediate impact on oil production. In other words, B is lower-triangular such that

$$\begin{bmatrix} U_t^{\Delta prod} \\ U_t^q \\ U_t^p \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{\text{oil-s}} \\ \varepsilon_t^{\text{aggr-d}} \\ \varepsilon_t^{\text{oil-d}} \end{bmatrix}. \quad (5)$$

Kilian and Murphy (2011) point out that a zero short-run oil supply elasticity of aggregate demand may not be realistic and the same is true for the short-run effect of an oil-market specific demand shock on real activity. Therefore Kilian and Murphy (2011) propose the use of sign restrictions that do not fix these effects at zero. More precisely, they impose the sign restrictions given in Table 1 for the impact effects. They also point out that these restrictions do not identify the shocks uniquely but result in a rather large set of admissible shocks and correspondingly imprecise impulse responses. They narrow down the range of admissible shocks by imposing bounds on the oil supply elasticity. In our approach we can get additional information from the data, as demonstrated in the following.

Table 1: Sign Restrictions for Impact Responses from Kilian and Murphy (2009)

	oil supply shock	aggregate demand shock	oil-market specific demand shock
oil production	–	+	+
real activity	–	+	–
oil price	+	+	+

Concerns regarding the exclusion restrictions imposed by Kilian (2009) are also expressed by Baumeister and Peersman (2010) who consider a similar system of variables. They use sign restrictions in the framework of a time-varying coefficient VAR model that allows for time-varying coefficients in addition to heteroskedasticity in the residuals. They find that during our sample period there was a rise in oil price volatility while at the same time the volatility of oil production has declined. Hence, it seems plausible to

use our approach and utilize the changes in the volatility of the shocks for identification.

3.2 The data and the VAR model

We use monthly data from 1973m2 - 2006m12, as in Kilian (2009). An updated dataset is used by Kilian and Murphy (2011). These authors use a VAR(24) model. However, model selection criteria favor smaller VAR orders. We base our analysis on a VAR(3) model that is selected by the more generous AIC. Notice that we have to use a nonlinear optimization algorithm for maximizing the likelihood function. Such algorithms run into problems for heavily parameterized models. Admittedly, standard residual tests indicate some remaining autocorrelation in the residuals of the VAR(3) model. The same is true for the VAR(24) model, however, and we have checked that qualitatively similar impulse responses result from both models in a conventional analysis.

In Figure 1 the residuals of both the VAR(24) and the VAR(3) models are plotted. While the two sets of residuals naturally look a bit differently, they both show that the residuals of oil production are much more volatile in the first part of the sample while the oil price residuals are substantially more volatile in the later sample period from the middle of the 1980s onwards. The figure hence confirms the observation made by Baumeister and Peersman (2010).

3.3 The number of MS states

In Table 2 some statistics are presented for a range of different models. In particular, models with different numbers of states and restrictions on the impact effects matrix B are compared. Considering only unrestricted VAR(3) models, the MS(3)-VAR(3) is favored by SC while AIC prefers a 4-state model. It is worth pointing out that models with MS are much preferred to the VAR(3) model without allowance for changes in volatility. Of course, this result is not surprising given the residual graphs in Figure 1. For the MS(3) models the SC and AIC values are further reduced by imposing the restriction of a state-invariant impact effects matrix B and lower-triangularity of this matrix. The SC value also declines when these restrictions are imposed and only MS(4) models are compared while the same is not true for AIC which is minimized for the least restricted model. Thus, based on the model selection criteria models with state invariant and recursive impact effects have some support although they are not favored generally.

In deciding on the number of MS states it may be worth looking at the smoothed state probabilities for the sample period. They are shown in Figure 2 where a state-invariant B matrix is imposed for the MS(3) and MS(4) models. The corresponding state covariance matrices are given in Table 3. Looking at the state probabilities of the MS(2)-VAR(3) model first, it can be seen that the first part of the sample is mainly associated with State 2 while State 1 dominates the second part of the sample. In Table 3 it can be seen that the volatility of oil production is relatively high in State 2 whereas the volatility of the oil price is much higher in State 1 than in State 2. This “puzzle” is further investigated by Baumeister and Peersman (2010), as mentioned earlier. For the purposes of our study the important point is that volatility changes occur during the sample period.

Looking at the smoothed probabilities of the MS(3) model in Figure 2 it becomes clear that the first two states are similar to those of the MS(2) model while the third state captures apparently some special periods. From Table 3 it becomes clear that the third state is characterized by very high volatility both in oil production and the real price of oil. Thus, it captures special states of particularly high volatility in the crude oil market. Many of the State 3 periods can be associated with specific events which were important for the crude oil market as listed in Barsky and Kilian (2004). For instance, there are peaks associated with the October War and oil embargo from October 1973 through early 1974, the Iranian Revolution between October 1978 and February 1979, the outbreak of the Iran-Iraq war in September 1980, the invasion of Kuwait in August 1990, and the important OPEC meeting in March 1999. Thus, the three-state model is not only favored by SC but also makes good sense when looking at the developments in the crude oil market.

The situation is somewhat different for the four-state model. The first and second states are associated with similar periods as in the MS(3) model, as seen in Figure 2 and also the associated covariance matrices in Table 3 are similar. Thus, the first state is again associated with periods of low volatility in oil production while the second state captures periods with higher volatility in the price of oil. In the third and fourth states the volatility of oil production is again very high while the volatility of the price of oil is very large in State 3 only. Unfortunately, State 3 is only associated with relatively few periods. Hence, the model is difficult to estimate and the estimates are unreliable. Overall, from a statistical point of view as well as considerations of events in the crude oil market, the model is more problematic than a MS(3) model and, hence, we favor the MS(3)-VAR(3) model in the following.

In Figure 3 the standardized residuals of the MS(3)-VAR(3) model are

presented. They are determined as

$$\hat{\Sigma}_{t|t-1}^{-1/2}\hat{u}_t,$$

where $\hat{u}_t = y_t - \hat{\nu} - \hat{A}_1 y_{t-1} - \dots - \hat{A}_p y_{t-p}$ and

$$\hat{\Sigma}_{t|t-1} = \sum_{m=1}^M \widehat{\Pr}(s_t = m | \mathbf{Y}_{t-1}) \hat{\Sigma}_m.$$

Here $\hat{\Sigma}_{t|t-1}$ denotes the estimated residual covariance matrix conditional on information up to time $t - 1$, that is, conditional on $\mathbf{Y}_{t-1} = (Y_{t-1}, \dots, Y_1)$. Clearly, the volatility of the standardized residuals in Figure 3 is more regular than in Figure 1. In other words, the MS structure captures the changes in volatility well. Hence, we will focus the analysis on the MS(3)-VAR(3) model in the following, which is also plausible from a subject matter point of view.

3.4 Statistical analysis of MS(3)-VAR(3) model

Because we are interested in using the MS structure for identification purposes, a main question of interest is whether a state-invariant initial effects matrix B is compatible with the data and whether the associated relative variances are sufficiently different to obtain a statistical identification of the shocks. In Table 2 we already see that both AIC and SC support a state-invariant B matrix for the MS(3) model. In Table 4 the likelihood ratio test from Lanne et al. (2010) for the null hypothesis of a state-invariant B is presented. Its p -value is 0.63 and hence exceeds conventional significance levels substantially. Thus, the LR test also supports a state-invariant initial effects matrix.

The estimated λ_{ij} s of the MS(3)-VAR(3) model with state-invariant B are shown in Table 5 together with estimated standard errors. The standard errors indicate that estimation precision is quite reasonable and, hence, we can hope for sufficient heterogeneity to get identification. Statistical tests of the relevant hypotheses are presented in Table 6. Recall that B is locally identified (identified apart from changes in sign and permutation of its columns) if for each pair of subscripts (i, j) there is a Λ_m matrix for which the corresponding diagonal elements are different, that is, $\lambda_{mi} \neq \lambda_{mj}$ for some $m \in \{2, \dots, M\}$. For our three-dimensional system we thus have to check three pairs of subscripts. The corresponding Wald and LR test results are shown in Table 6. The Wald tests are attractive in the present context because they are very easy to compute from the estimates of the model with state-invariant B . On the other hand, they are known to be unreliable for

highly nonlinear null hypotheses. Two of the three null hypotheses are clearly rejected by the Wald tests at a 5% significance level while the third one is rejected at a 10% level.

We also computed computationally more demanding LR tests. They require an additional likelihood optimization under the null hypothesis, which is a challenging task for the presently considered model class with the risk of ending up in some local optimum. Note that the LR tests are computed for models where the diagonal elements of Λ_2 are sorted in increasing order. The LR test values result in very small p -values and strongly support identification of B . Thus, we conclude that the data are in line with identified shocks. These shocks are the only ones in our framework that are time-invariant with the same impact effects across states. Note that time-invariant shocks and impulse responses are also assumed by Kilian (2009) and Kilian and Murphy (2011), but not by Baumeister and Peersman (2010), who allow for time-varying coefficients. We emphasize that our identifying restrictions so far are based on data driven procedures and the data are not objecting to them in the framework of the MS(3)-VAR(3) model.

Of course, our identification is a statistical one and there is still a question whether it is consistent with economically meaningful shocks, in particular with the shocks considered by Kilian (2009) and Kilian and Murphy (2011). The question whether our statistically identified shocks are in line with the sign identified shocks considered by Kilian and Murphy (2011) will be considered when we look at impulse responses in the next subsection. In the present context we can, however, perform a statistical test of the exclusion restrictions used by Kilian (2009), which are given in Equation (5). Having identifying information for the shocks from other sources, allows us to apply simple LR tests because Kilian's recursive identification scheme becomes over-identifying in the context of our model, as explained in Section 2. The AIC and SC values for models with state-invariant, lower-triangular B matrix in Table 2 and the corresponding LR tests for a MS(3)-VAR(3) model presented in Table 4 all support a lower-triangular B . In particular, the LR tests in Table 4 have very large p -values and, hence, do not reject lower-triangularity of B . This holds for both a test against an unrestricted MS(3) model and one against a MS(3) model with state-invariant B .

Further support for a lower-triangular B is presented in Table 5, where the estimated λ_{ij} s of the model with a lower-triangular B matrix are given. They are very close to those of the corresponding model with unrestricted B . Note that the columns of B can no longer be permuted when the matrix is triangular. Thus, the ordering of the λ_{ij} s is the one corresponding to the lower-triangular B matrix. It results in λ_{2j} s in decreasing order and, hence, we order the λ_{2j} s for the model with unrestricted, state-invariant B

accordingly. Apparently, from a statistical point of view, the assumption of a lower-triangular initial effects matrix is not restrictive. In the next section it will be explored whether the sign restrictions used in the related literature are also in line with our statistically identified shocks. An impulse response analysis will be performed for that purpose.

3.5 Impulse response analysis

Given that we have identified shocks, we can also compute impulse responses from our MS(3)-VAR(3) model with state-invariant initial effects matrix B . These are displayed in Figure 4 together with 68% bootstrap confidence intervals. The scaling of the impulse responses is determined by B , which, in turn, is scaled such that the residuals in State 1 have identity covariance matrix. This is to some extent arbitrary because the numbering of the states is arbitrary. The scaling is also quite different from that used by Kilian. We emphasize that our normalization of the shocks in the first state affects only the scaling but not the shape and the sign of the impulse responses.

A priori we do not know which economically interpretable shocks are represented by the statistically identified shocks. However, the fact that our ordering of the shocks is in line with a lower-triangular, recursive identification scheme indicates that they may be labeled as in (5), that is, the first shock may be viewed as an oil supply shock, the second as an aggregate demand shock and the third as an oil-market specific demand shock. With this labeling in mind, they can be seen to be in line with the sign restrictions given in Table 1. Only the oil-market specific demand shock could be viewed as potentially problematic because its initial effect on output tends to be positive. Zero is really at the lower end of the 68% confidence interval. Because the confidence level could be regarded as pretty low, our analysis provides little basis for questioning the sign restrictions. Using a slightly larger confidence level would result in intervals with negative values for the initial response of output. In other words, the data do not object to the sign restrictions, that is, the system as set up here and the data used are consistent with the sign restrictions adopted by Kilian and Murphy (2011).

Notice that the impulse responses of oil production are cumulated responses of $\Delta prod_t$. They fall outside the respective 68% bootstrap confidence bands. That feature is also observed in other studies and is not uncommon in the SVAR literature. It may be due to the skewness and biasedness of the distribution of the impulse response estimates.

The fact that our method does not exclude any of the restrictions imposed in the previous studies may be seen as a weakness indicating lack of power of our procedure. In defense of our method we note that the restrictions in

Table 1 are indeed not very tight. In the literature using sign restrictions it is not common to restrict only the impact effects. Assuming that the sign restrictions hold for some quarters after the shock has occurred would be a more common set of restrictions in the related literature. If we consider such restrictions for illustrative purposes, it turns out that our shocks do not satisfy them. For example, if we stick with the 68% confidence intervals around the impulse responses and consider this as the possible range of impulse responses supported by the data, there is no shock that raises oil production, reduces real activity and increases the prize of oil for four quarters after hitting the system. Thus, in our framework the data do not support the existence of an oil-market specific demand shock satisfying such more stringent conditions. This result illustrates that our procedure does have some discriminatory power. The restrictions used by Kilian and Murphy (2011) are apparently very weak and, hence, difficult to reject by the data, although more stringent restrictions can be rejected. Of course, imposing weak restrictions and still being able to draw conclusions from the analysis is an advantage of the Kilian-Murphy study and our analysis gives further support to their results.

Kilian and Murphy (2011) show that if only the sign restrictions from Table 1 are used, some elasticities become unreasonably large. Therefore they search for other types of information that can be used to obtain more reasonable responses. As mentioned earlier, they impose also restrictions on the oil supply elasticity. In the context of our approach we can use the statistical identification restrictions and obtain the impulse responses with 68% confidence intervals in Figure 4. In some cases they are also not very precise but can be restricted further by imposing statistically valid restrictions such as lower-triangularity of B . The corresponding impulse responses are depicted in Figure 5. Overall they are not much more precise than those from the unrestricted model in Figure 4, apart from the impact effects of aggregate demand and oil-specific demand shocks that are now restricted to zero.

One may of course argue that the main gains in precision relative to the impulse responses shown in Kilian and Murphy (2011) and Kilian (2009) come from the smaller VAR lag order. Recall that these authors use lag order $p = 24$ while we use $p = 3$. Indeed our smaller order results in smoother impulse responses. They are qualitatively similar to those from Kilian (2009), however. To show this, we present the impulse responses of a conventional, recursively identified VAR(24) in Figure 6 and the corresponding quantities from a VAR(3) in Figure 7. The confidence intervals shown in these figures are obtained from a standard percentile procedure. The ones in Figure 6 look very similar to the corresponding ones in Figure 3 of Kilian (2009). Notice that Kilian reports one- and two-standard error bounds that are naturally

symmetric around the impulse responses while we use (potentially asymmetric) 68% bounds for ease of comparison with our previous results.

The overall conclusion from our study is that using the statistical information more fully allows for more precise inference on the effects of structural shocks in a SVAR analysis. It provides additional identifying information that can complement the information from economic theory and allows to check restrictions that cannot be checked against the data in a conventional analysis with exclusion or sign restrictions. It also has the advantage of providing information on the overall compatibility of the model and the data used with the theoretical considerations of economists. In other words, it can provide information on the compatibility of the data despite potential deviations from an ideal model world due to omitted variables, measurement errors or other data problems.

4 Conclusions

We argue that restrictions derived from some economic model may not be valid in a corresponding empirical model. Reasons may be that the economic model is a simplification of reality that explicitly ignores important effects in the real economy. For example, there may be an omitted variables problem if only those variables are included in the empirical model that are described in the theoretical model. Moreover, the empirical variables may not correspond precisely to the variables underlying the economic model due to measurement errors, data adjustments such as trend or seasonal adjustments or because the theoretical variables are not measurable. In a conventional SVAR analysis one would still impose the restrictions suggested by economic reasoning. If the restrictions just-identify the shocks it may go unnoticed that the empirical model does not reflect the actual relationships between the variables involved. Similarly, if only sign restrictions are used to identify the shocks, in a conventional analysis the data have no basis to speak up against the model. Therefore there are proposals to use statistical properties such as heteroskedasticity to complement economic restrictions and check restrictions that are just-identifying in a conventional SVAR framework. In this study we propose to use similar devices for checking sign restrictions.

More precisely, we propose to model changes in the volatility of the shocks by a MS mechanism and we show how this device can be used to check sign restrictions in addition to exclusion restrictions on the impact and long-run effects of the shocks. If the volatility structure is rich enough it may provide identified shocks which should correspond to the economic shocks if the empirical model is in line with the economic model. In this case, the

statistically identified shocks have to satisfy, for example, the sign restrictions or other restrictions derived from economic considerations.

For illustrative purposes we use a small, three-dimensional model of the market for crude oil consisting of the change in oil production, the price of oil and an index of economic activity. For this model, oil supply and demand shocks as well as an aggregate demand shock, are specified with sign restrictions for the corresponding impulse responses, on the one hand, and changes in the volatility of the residuals are observed on the other. The changes in volatility are used to obtain identified shocks. It turns out that these shocks satisfy a previously used weak set of sign restrictions but not a slightly stricter set of such restrictions. Our analysis allows us not only to check previously used sign restrictions but it also enables us to impose further exclusion restrictions that may help obtain more precisely estimated impulse responses. More generally, the study demonstrates how identifying restrictions in general and sign restrictions in particular can be checked in our framework.

References

- Amisano, G. and Giannini, C. (1997). *Topics in Structural VAR Econometrics*, 2nd edn, Springer, Berlin.
- Barsky, R. B. and Kilian, L. (2004). Oil and the macroeconomy since the 1970s, *Journal of Economic Perspectives* **18**: 115–134.
- Baumeister, C. and Peersman, G. (2010). Sources of the volatility puzzle in the crude oil market, *Working Paper 2010/634*, Universiteit Gent.
- Blanchard, O. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances, *American Economic Review* **79**: 655–673.
- Bouakez, H. and Normandin, M. (2010). Fluctuations in the foreign exchange market: How important are monetary policy shocks?, *Journal of International Economics* **81**: 139153.
- Canova, F. (2002). Validating monetary DSGE models through VARs, *Discussion Paper 3442*, CEPR.
- Canova, F. and De Nicoló, G. (2002). Monetary disturbances matter for business fluctuations in the G-7, *Journal of Monetary Economics* **49**: 1131–1159.
- Canova, F. and Pappa, E. (2007). Price differentials in monetary unions: The role of fiscal shocks, *Economic Journal* **117**: 713–734.

- Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII US, *Review of Economic Dynamics* **8**: 262–302.
- Dedola, L. and Neri, S. (2007). What does a technology shock do? A VAR analysis with model-based sign restrictions, *Journal of Monetary Economics* **54**: 512–549.
- Faust, J. (1998). The robustness of identified VAR conclusions about money, *Carnegie-Rochester Conference Series in Public Policy* **49**: 207–244.
- Herwartz, H. and Lütkepohl, H. (2011). Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks, *Working Paper series*, EUI, Florence.
- Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market, *American Economic Review* **99**: 1053–1069.
- Kilian, L. and Murphy, D. (2011). Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market VAR models, *Journal of the European Economic Association* .
- King, R. G., Plosser, C. I., Stock, J. H. and Watson, M. W. (1991). Stochastic trends and economic fluctuations, *American Economic Review* **81**: 819–840.
- Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*, Springer-Verlag, Berlin.
- Lanne, M. and Lütkepohl, H. (2008). Identifying monetary policy shocks via changes in volatility, *Journal of Money, Credit and Banking* **40**: 1131–1149.
- Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). Structural vector autoregressions with Markov switching, *Journal of Economic Dynamics and Control* **34**: 121–131.
- Lippi, F. and Nobili, A. (2011). Oil and the macroeconomy: A quantitative structural analysis, *Journal of the European Economic Association* .
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.
- Moon, H. R. and Schorfheide, F. (2009). Bayesian and frequentist inference in partially identified models, *Working Paper 14882*, NBER.
- Moon, H. R., Schorfheide, F., Granziera, E. and Lee, M. (2009). Inference for VARs identified with sign restrictions, *manuscript*, USC, Los Angeles.

- Mountford, A. and Uhlig, H. (2009). What are the effects of fiscal policy shocks?, *Journal of Applied Econometrics* **24**: 960–992.
- Normandin, M. and Phaneuf, L. (2004). Monetary policy shocks: Testing identification conditions under time-varying conditional volatility, *Journal of Monetary Economics* **51**: 1217–1243.
- Pappa, E. (2009). The effects of fiscal shocks on employment and the real wage, *International Economic Review* **50**: 217–244.
- Peersman, G. and Straub, R. (2009). Technology shocks and robust sign restrictions in a euro area SVAR, *International Economic Review* **50**: 727–750.
- Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy, *Review of Economic Studies* **72**: 821–852.
- Rigobon, R. (2003). Identification through heteroskedasticity, *Review of Economics and Statistics* **85**: 777–792.
- Rigobon, R. and Sack, B. (2003). Measuring the reaction of monetary policy to the stock market, *Quarterly Journal of Economics* **118**: 639–669.
- Rubio-Ramirez, J. F., Waggoner, D. and Zha, T. (2005). Markov-switching structural vector autoregressions: Theory and applications, *Discussion Paper*, Federal Reserve Bank of Atlanta.
- Sims, C. A. (1980). Macroeconomics and reality, *Econometrica* **48**: 1–48.
- Sims, C. A., Waggoner, D. F. and Zha, T. (2008). Methods for inference in large multiple-equation Markov-switching models, *Journal of Econometrics* **146**: 255–274.
- Sims, C. A. and Zha, T. (2006). Were there regime switches in U.S. monetary policy?, *American Economic Review* **96**: 54–81.
- Uhlig, H. (2005). What are the effects of monetary policy on output? Results from an agnostic identification procedure, *Journal of Monetary Economics* **52**: 381–419.

Table 2: Comparison of MS-VAR(3) Models for $Y_t = (\Delta prod_t, q_t, p_t)'$

Model	$\log L_T$	AIC	SC
VAR(3) without MS	-4271.47	8614.94	8758.99
MS(2), unrestricted	-4106.82	8301.64	8477.71
MS(2), B lower-triangular	-4108.85	8299.70	8463.76
MS(3), unrestricted	-4048.84	8205.68	8421.76
MS(3), state-invariant B	-4049.70	8201.40	8405.47
MS(3), state-invariant, lower-triangular B	-4050.13	8196.27	8388.34
MS(4), unrestricted	-4013.32	8158.64	8422.73
MS(4), state-invariant B	-4030.42	8180.85	8420.93
MS(4), state-invariant, lower-triangular B	-4038.40	8190.81	8418.89

Note: L_T - likelihood function, $AIC = -2\log L_T + 2 \times \text{no of free parameters}$, $SC = -2\log L_T + \log T \times \text{no of free parameters}$.

Table 3: Estimated State Covariance Matrices of MS(m)-VAR(3) Models with State-Invariant B , $m = 2, 3, 4$, for $Y_t = (\Delta prod_t, q_t, p_t)'$

	$m = 2$	$m = 3$	$m = 4$
Σ_1	$\begin{bmatrix} 155.59 & & & \\ 2.02 & 12.38 & & \\ -9.43 & 2.36 & 58.29 & \end{bmatrix}$	$\begin{bmatrix} 99.38 & & & \\ 0.96 & 10.51 & & \\ -0.86 & 1.09 & 42.84 & \end{bmatrix}$	$\begin{bmatrix} 90.92 & & & \\ 1.33 & 9.89 & & \\ -0.98 & 0.49 & 40.27 & \end{bmatrix}$
Σ_2	$\begin{bmatrix} 912.86 & & & \\ 18.49 & 35.38 & & \\ -0.56 & 0.56 & 2.94 & \end{bmatrix}$	$\begin{bmatrix} 533.95 & & & \\ 9.39 & 33.86 & & \\ 5.07 & 0.64 & 2.45 & \end{bmatrix}$	$\begin{bmatrix} 368.56 & & & \\ 6.42 & 23.89 & & \\ -0.60 & 1.06 & 1.52 & \end{bmatrix}$
Σ_3		$\begin{bmatrix} 2499.03 & & & \\ 65.38 & 37.42 & & \\ 16.46 & 5.31 & 190.49 & \end{bmatrix}$	$\begin{bmatrix} 947.64 & & & \\ 17.89 & 39.83 & & \\ -5.97 & 2.03 & 222.26 & \end{bmatrix}$
Σ_4			$\begin{bmatrix} 2365.54 & & & \\ 47.36 & 57.70 & & \\ -3.55 & 2.49 & 9.59 & \end{bmatrix}$

Table 4: LR Tests of Restrictions for MS(3)-VAR(3) Models for $Y_t = (\Delta prod_t, q_t, p_t)'$

H_0	H_1	LR	df	p -value
state-invariant B	unrestricted MS(3)-VAR(3)	1.72	3	0.63
state-invariant, lower-triangular B	state-invariant B	0.86	3	0.84
state-invariant, lower-triangular B	unrestricted MS(3)-VAR(3)	2.58	6	0.86

Note: $LR = 2(\log L_T - \log L_T^r)$, where L_T^r denotes the maximum likelihood under H_0 and L_T denotes the maximum likelihood for the model under H_1 .

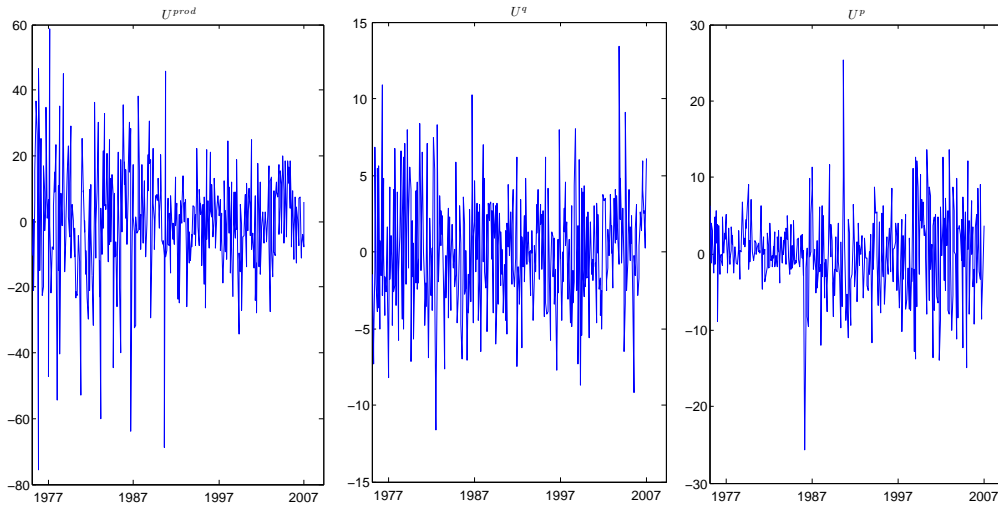
Table 5: Estimates of Structural Parameters of MS(3)-VAR(3) Models for $Y_t = (\Delta prod_t, q_t, p_t)'$ with State-invariant B

parameter	unrestricted B		lower-triangular B	
	estimate	std.dev.	estimate	std.dev.
λ_{21}	5.384	0.933	5.429	0.944
λ_{22}	3.210	0.485	3.220	0.489
λ_{23}	0.056	0.009	0.057	0.009
λ_{31}	25.235	6.938	25.097	7.107
λ_{32}	3.387	1.242	3.481	1.258
λ_{33}	4.441	1.018	4.382	1.011

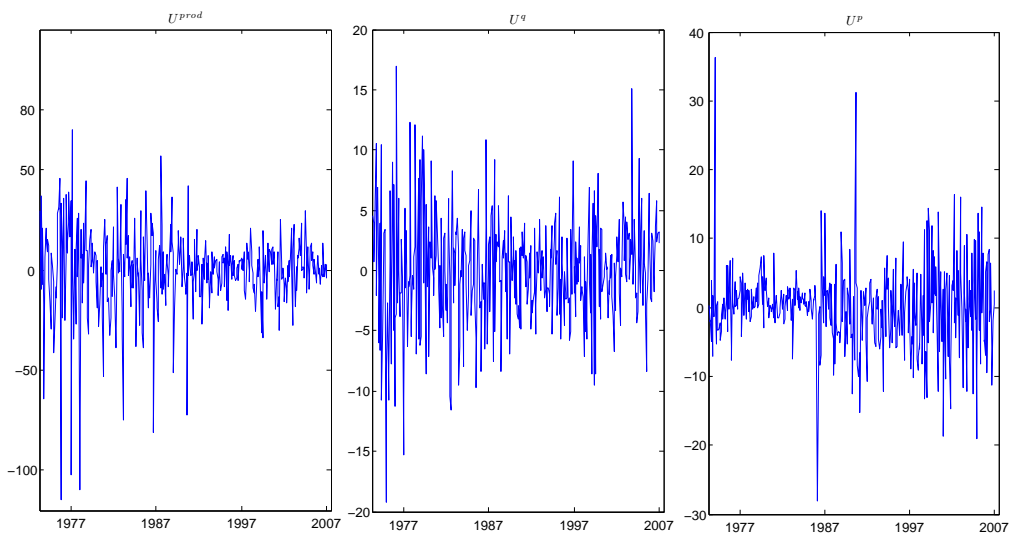
Note: Standard errors are obtained from the inverse of the outer product of numerical first order derivatives.

Table 6: Tests for Equality of λ_{ijs} for MS(3)-VAR(3) Model with State-invariant B

H_0	Wald statistic	p -value	LR statistic	p -value
$\lambda_{21} = \lambda_{22}, \lambda_{31} = \lambda_{32}$	7.99	0.02	20.39	3.7×10^{-5}
$\lambda_{21} = \lambda_{23}, \lambda_{31} = \lambda_{33}$	7.87	0.02	21.04	2.7×10^{-5}
$\lambda_{22} = \lambda_{23}, \lambda_{32} = \lambda_{33}$	5.16	0.07	27.15	1.3×10^{-6}

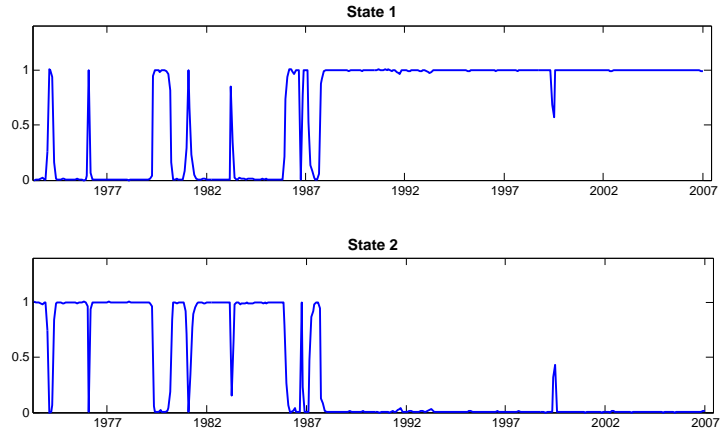


(a) Residuals of VAR(24) model

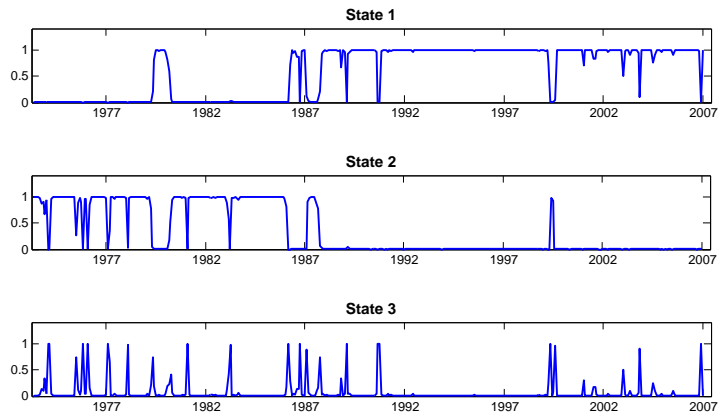


(b) Residuals of VAR(3) model

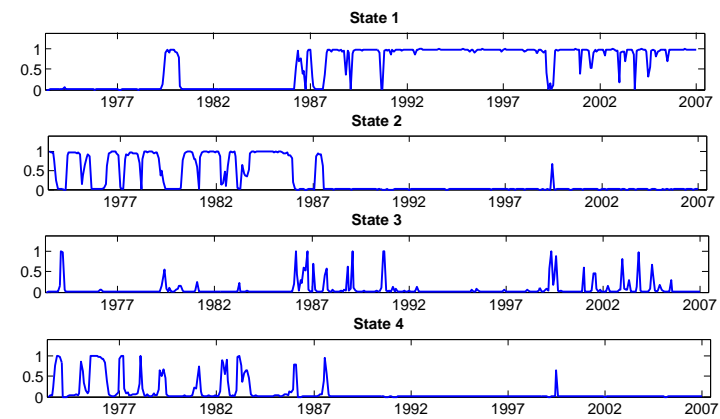
Figure 1: Residuals of VAR(24) and VAR(3) models.



(a) Smoothed state probabilities for MS(2)-VAR(3)



(b) Smoothed state probabilities for MS(3)-VAR(3)



(c) Smoothed state probabilities for MS(4)-VAR(3)

Figure 2: Smoothed state probabilities of unrestricted MS(m)-VAR(3) models for $m = 2, 3, 4$.

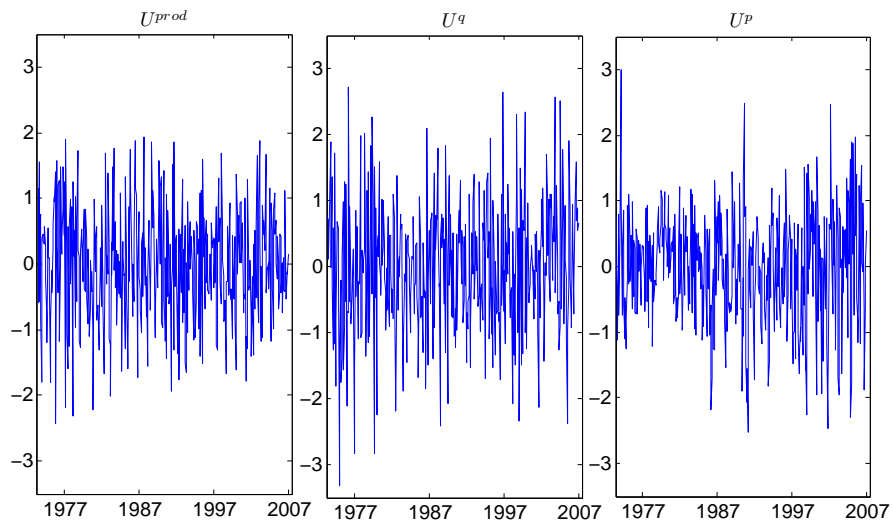


Figure 3: Standardized residuals of MS(3)-VAR(3) model with state-invariant B .

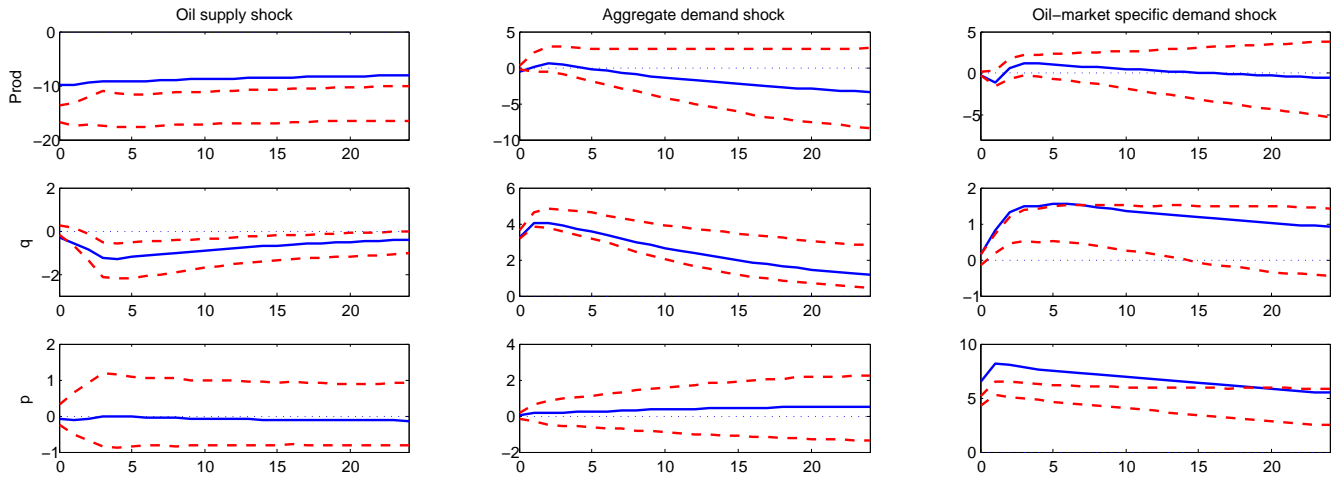


Figure 4: Impulse responses with 68% confidence bounds of the MS(3)-VAR(3) model with state-invariant B .

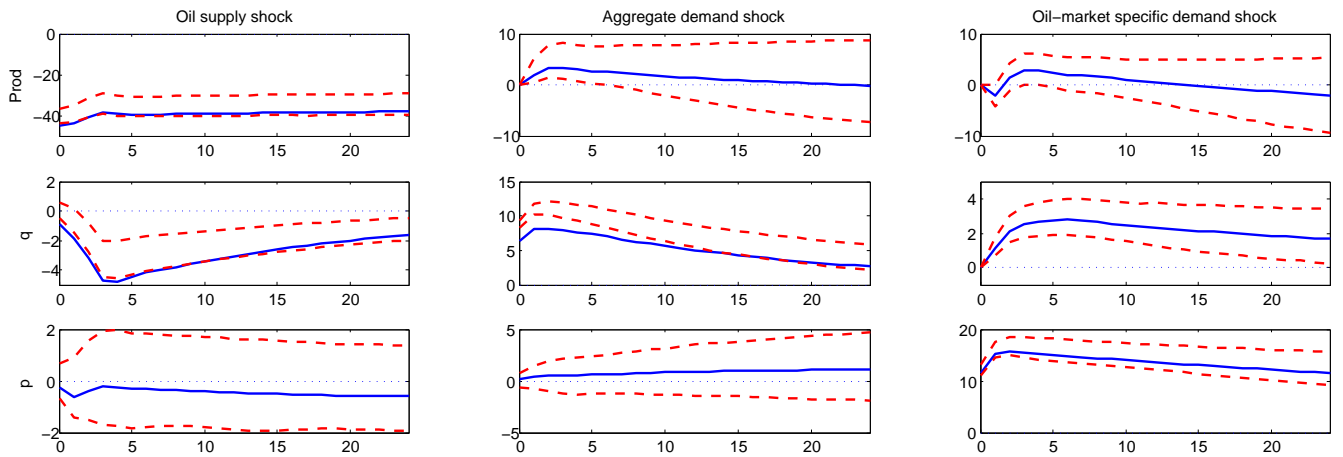


Figure 5: Impulse responses with 68% confidence bounds of the MS(3)-VAR(3) model with state-invariant, lower-triangular B .

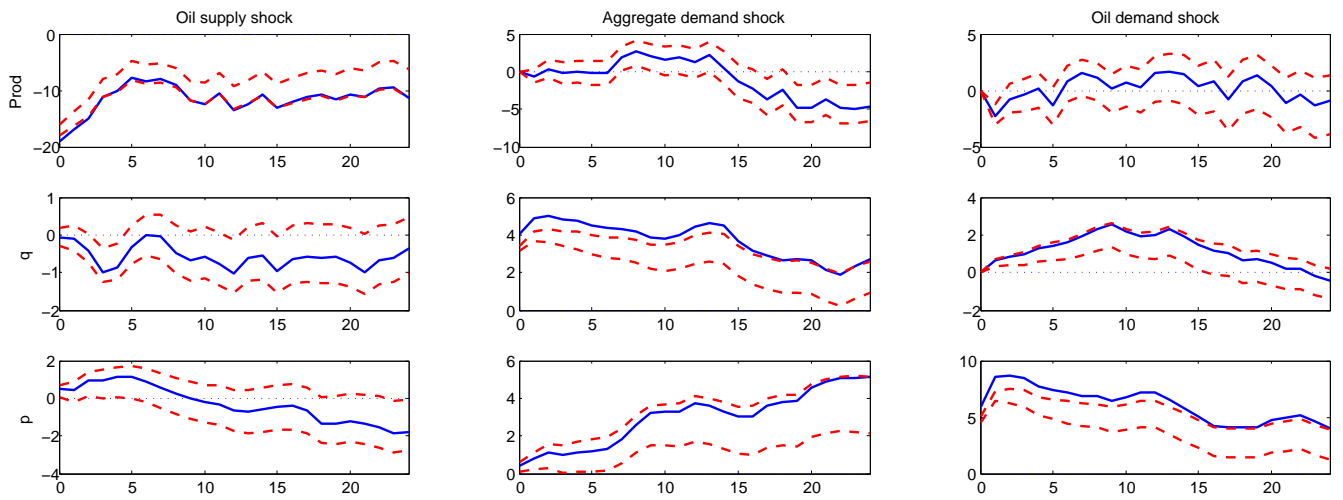


Figure 6: Impulse responses with 68% confidence bounds of a recursively identified VAR(24) model.

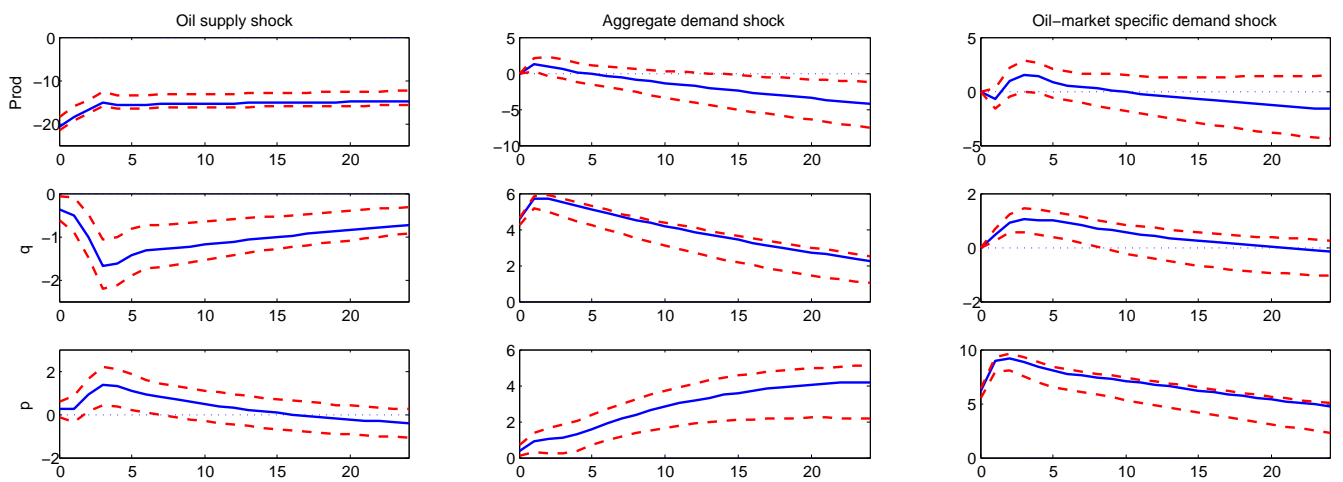


Figure 7: Impulse responses with 68% confidence bounds of a recursively identified VAR(3) model.